NORTH-HOLLAND

On Error Bounds for Eigenvalues of a Matrix Pencil

Mario Ahues
Équipe d'Analyse Numérique UMR 5585 CNRS
Université Jean Monnet
23, rue Dr. Paul Michelon
42023 Saint-Étienne, France

and

Balmohan Limaye
Department of Mathematics
Indian Institute of Technology
Powai, Bombay 400 076, India

Submitted by Richard A. Brualdi

ABSTRACT

An error bound for approximate eigenvalues of a complex n-dimensional pencil (A, B) is given. From our theorem several well-known bounds follow as corollaries. Our result takes into account the general residual $AX - BXW$, where $X \in \mathbb{C}^{n\times m}$ and $W \in \mathbb{C}^{m\times m}$ with $m \leqslant n$. © 1998 Elsevier Science Inc.

1. PRELIMINARIES AND NOTATION

Let n, m be integers such that $n \geqslant m \geqslant 1$. $\mathbb{C}^n$ will denote the complex linear space of column vectors with n complex coordinates, $\mathbb{C}^{n\times m}$ the space of $n \times m$ complex matrices, $\mathscr{L}(\mathbb{C}^{n\times m})$ the space of linear operators from $\mathbb{C}^{n\times m}$ into itself. Capital letters A, B, X, W, etc. will be used to denote matrices with more than one column. I denotes any identity matrix. Small letters such as x, y will denote elements of $\mathbb{C}^n$ with $n > 1$. Bold capital letters such as $\mathbf{T}$ will denote elements of $\mathscr{L}(\mathbb{C}^{n\times m})$. $\mathbf{I}$ will denote the identity operator in this space.

LINEAR ALGEBRA AND ITS APPLICATIONS 268:71–89 (1998)

0024-3795/98/$19.00
PII S0024-3795(97)000028-1

Let $A \in \mathbb{C}^{n \times n}$ and $B \in \mathbb{C}^{n \times n}$ be such that the pencil (A, B) is regular, that is, there exists $\lambda \in \mathbb{C}$ such that $\det(A - \lambda B)$ is not zero. Consider a matrix $W \in \mathbb{C}^{m \times m}$.

We remark that if $X \in \mathbb{C}^{n \times m}$ has m linearly independent columns (for instance, if $X^*X = I$), and if $R \equiv AX - BXW = 0$, then any eigenvalue of W is an eigenvalue of the pencil (A, B). The object of this paper is to obtain bounds between the eigenvalues of W and those of (A, B) when R is small.

If $T : \mathbb{C}^{n \times m} \to \mathbb{C}^{n \times m}$ is defined by $\mathbf{T}(X) = AX - BXW$ and $\mathbf{B} : \mathbb{C}^{n \times m} \to \mathbb{C}^{n \times m}$ by $\mathbf{B}(X) = BX$, then the eigenvalues of the pencil $(\mathbf{T}, \mathbf{B})$ are the differences between the eigenvalues of (A, B) and W. When $B = I$, this result is known as the Frobenius theorem.

Given $x_1, \dots, x_m \in \mathbb{C}^n$, let

$$X = [x_1, \dots, x_m] \in \mathbb{C}^{n \times m} \quad \text{and} \quad \operatorname{vec}(X) = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \in \mathbb{C}^{nm}.$$

Given a linear map $\mathbf{T} : \mathbb{C}^{n \times m} \to \mathbb{C}^{n \times m}$, we obtain a linear map $\mathscr{T} : \mathbb{C}^{nm} \to \mathbb{C}^{nm}$ as follows: For $x \in \mathbb{C}^{nm}$, there exists a unique $X \in \mathbb{C}^{n \times m}$ such that $x = \operatorname{vec}(X)$. We define

$$\mathscr{T}x = \operatorname{vec}(\mathbf{T}(X)).$$

It is clear that the matrix representing $\mathbf{T}$ in the canonical basis of $\mathbb{C}^{n \times m}$ is the same as the matrix representing $\mathscr{T}$ in the canonical basis of $\mathbb{C}^{nm}$.

Let $\| \ \|_2$ denote the Euclidean norm on $\mathbb{C}^n$, and $\| \ \|_F$ denote the Frobenius norm on $\mathbb{C}^{n \times m}$;

$$\left\| \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \right\|_2 = \left(\sum_{i=1}^{n} |\xi_i|^2 \right)^{1/2},$$

$$\left\| \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1m} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nm} \end{bmatrix} \right\|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} |\alpha_{ij}|^2 \right)^{1/2}.$$

Clearly, $\|X\|_F = \|\operatorname{vec}(X)\|_2$. The subordinated operator norms will also be denoted by $\| \ \|_2$ and $\| \ \|_F$, as the case may be.

The results that follow will be needed in the sequel.

PROPERTY 1.1. *If* $X \in \mathbb{C}^{n\times m}$, $P \in \mathbb{C}^{n\times n}$, *and* $Q \in \mathbb{C}^{m\times m}$, *then*

$$\|PX\|_F \leqslant \|P\|_2\|X\|_F \quad \textit{and} \quad \|XQ\|_F \leqslant \|X\|_F\|Q\|_2.$$

Proof. Theorem II 3.9 of [18]. ∎

PROPERTY 1.2. *If* $\mathbf{T} \in \mathscr{L}(\mathbb{C}^{n\times m})$, *then* $\|\mathbf{T}\|_F = \|\mathscr{T}\|_2$.

Proof.

$$\begin{aligned}\|\mathbf{T}\|_F &= \sup\{\|\mathbf{T}(X)\|_F : X \in \mathbb{C}^{n\times m}, \|X\|_F = 1\}\\ &= \sup\{\|\operatorname{vec}(\mathbf{T}(X))\|_2 : X \in \mathbb{C}^{n\times m}, \|\operatorname{vec}(X)\|_2 = 1\}\\ &= \sup\{\|\mathscr{T}x\|_2 : x \in \mathbb{C}^{nm}, \|x\|_2 = 1\}\\ &= \|\mathscr{T}\|_2.\end{aligned}$$ ∎

PROPERTY 1.3. *Let* $\mathbf{T}, \mathbf{S} \in \mathscr{L}(\mathbb{C}^{n\times m})$ *and* $Q \in \mathbb{C}^{m\times m}$ *be such that* $Q^*Q = I$ *and* $\mathbf{T}(X) = [\mathbf{S}(XQ)]Q^*$ *for* $X \in \mathbb{C}^{n\times m}$. *Then* $\|\mathbf{T}\|_F = \|\mathbf{S}\|_F$.

Proof. It is easy to see that $\|\mathbf{T}\|_F \leqslant \|\mathbf{S}\|_F$. Interchanging the roles of $\mathbf{T}$ and $\mathbf{S}$ as well as those of Q and Q^*, we obtain $\|\mathbf{S}\|_F \leqslant \|\mathbf{T}\|_F$. ∎

Let $\operatorname{sp}(A, B)$ denote the generalized spectrum of the pencil (A, B). If B is invertible, we have

$$\operatorname{sp}(A, B) = \{\lambda \in \mathbb{C} : \det(A - \lambda B) = 0\}.$$

Also, let $\operatorname{sp}(W) = \operatorname{sp}(W, I)$ denote the (ordinary) spectrum of W.

We consider the linear operator $\mathbf{T} : \mathbb{C}^{n\times m} \to \mathbb{C}^{n\times m}$ defined by

$$\mathbf{T}(X) = AX - BXW \qquad \text{for} \quad X \in \mathbb{C}^{n\times m}.$$

PROPERTY 1.4. *The matrix* T *representing* $\mathbf{T}$ *in the canonical basis of* $\mathbb{C}^{n\times m}$ *is*

$$T = \begin{bmatrix} A - w_{11}B & -w_{21}B & \cdots & -w_{m1}B \\ -w_{12}B & A - w_{22}B & \cdots & -w_{m2}B \\ \vdots & & \ddots & \vdots \\ -w_{1m}B & -w_{2m}B & \cdots & A - w_{mm}B \end{bmatrix}.$$

Proof. Cf. [11]. ∎

It may be noted that if A and B are upper triangular and W is lower triangular, then T is upper triangular. This key observation is utilized to obtain bounds for the generalized eigenvalue problem. In particular, when $m = 1$ and $B = I$, bounds for the ordinary eigenvalue problem follow. Even in this case, some of the earlier results will be improved.

PROPERTY 1.5. *There exist invertible $n \times n$ complex matrices U and V, diagonal $n \times n$ complex matrices*

$$D_A = \operatorname{diag}(\alpha_1, \ldots, \alpha_n) \quad \text{and} \quad D_B = \operatorname{diag}(\beta_1, \ldots, \beta_n),$$

and strictly upper triangular $n \times n$ complex matrices N_A and N_B such that

$$U^{-1}AV = D_A + N_A \quad \text{and} \quad U^{-1}BV = D_B + N_B.$$

Moreover, if B is invertible, then

$$\operatorname{sp}(A, B) = \operatorname{sp}(D_A, D_B) = \bigcup_{i=1}^{n} \left\{ \frac{\alpha_i}{\beta_i} \right\}.$$

If B is Hermitian positive definite, then the matrices U and V may be chosen such that

$$U^{-1} = V^* \quad \text{and} \quad V^*BV = I.$$

Proof. See Theorems VI 1.9 and VI 1.16 of [18]. ∎

We fix U and V as given above. Although one may also require that U and V be unitary, we do not do so in general.

Schur's lemma implies that there exists a unitary matrix $Q \in \mathbb{C}^{m \times m}$ such that

$$L = (l_{ij}) = Q^*WQ$$

is lower triangular. Let us define the operators $\mathbf{D}$, $\tilde{\mathbf{D}}$, $\mathbf{N}$, and $\tilde{\mathbf{N}}$ belonging to $\mathscr{L}(\mathbb{C}^{n \times m})$ by

$$\mathbf{D}(X) = D_A X - D_B XW,$$

$$\tilde{\mathbf{D}}(X) = D_A X - D_B XL,$$

$$\mathbf{N}(X) = N_A X - N_B XW,$$

$$\tilde{\mathbf{N}}(X) = N_A X - N_B XL.$$

PROPERTY 1.6. *The following assertions are equivalent*:

(i) $\mathbf{T}$ *is invertible*,
(ii) $\mathbf{D}$ *is invertible*,
(iii) $\tilde{\mathbf{D}}$ *is invertible*,
(iv) $\mathrm{sp}(W) \cap \mathrm{sp}(A, B) = \varnothing$.

Proof. The equivalence of (i), (ii), and (iii) is easy. The equivalence with (iv) may be found in Theorem IV 2.1 of [9]. ∎

PROPERTY 1.7. *If* $\mathbf{T}$ *is invertible, then*

(i) $\mathbf{D}^{-1}(X) = (\tilde{\mathbf{D}}^{-1}(XQ))Q^*$ *for all* $X \in \mathbb{C}^{n\times m}$,
(ii) $\|\mathbf{D}^{-1}\|_F = \|\tilde{\mathbf{D}}^{-1}\|_F$,
(iii) $V[(\mathbf{I} + \mathbf{D}^{-1}\mathbf{N})^{-1}\mathbf{D}^{-1}][U^{-1}\mathbf{T}(X)] = X$ *for all* $X \in \mathbb{C}^{n\times m}$, *and*
(iv) *there exists a positive integer* l *such that* $(\mathbf{D}^{-1}\mathbf{N})^l = \mathbf{O}$.

Proof. (i): It is easy to check that for all $Y \in \mathbb{C}^{n\times m}$, $\mathbf{D}(Y) = [\tilde{\mathbf{D}}(YQ)]Q^*$. Setting $Y = \mathbf{D}^{-1}(X)$ and $\tilde{Y} = [\tilde{\mathbf{D}}^{-1}(XQ)]Q^*$, one obtains $[\tilde{\mathbf{D}}(\tilde{Y}Q)]Q^* = X = \mathbf{D}(Y) = [\mathbf{D}(YQ)]Q^*$, so that $\tilde{Y} = Y$.

(ii): This follows from Property 1.3.

(iii): For each $X \in \mathbb{C}^{n\times m}$,

$$\begin{aligned} &V\left[(\mathbf{I} + \mathbf{D}^{-1}\mathbf{N})^{-1}\mathbf{D}^{-1}\right]\left[U^{-1}(AX - BXW)\right] \\ &\quad = V\left[(\mathbf{I} + \mathbf{D}^{-1}\mathbf{N})^{-1}\mathbf{D}^{-1}(\mathbf{D} + \mathbf{N})\right](V^{-1}X) \\ &\quad = X. \end{aligned}$$

(iv): The matrix representing $\mathbf{D}^{-1}\mathbf{N}$ in the canonical basis of $\mathbb{C}^{n\times m}$ is strictly upper triangular. ∎

2. THE MAIN RESULT

In what follows we shall assume that B is invertible. Associated with L and D_B we define

$$\gamma = \begin{cases} 0 & \text{if} \quad m = 1, \\ \max\{|l_{ij}| : 1 \leqslant j < i \leqslant m\} & \text{if} \quad m > 1, \end{cases}$$

$$\bar{\beta} = \max\{|\beta_i| : 1 \leqslant i \leqslant n\},$$

$$\underline{\beta} = \min\{|\beta_i| : 1 \leqslant i \leqslant n\}.$$

We use the notation $\mathbb{R}_+ = \{t \in \mathbb{R} : t \geqslant 0\}$.
Let $\varphi_\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ be the function defined by

$$\varphi_\gamma(t) = \begin{cases} t^{m-1}\sqrt{m} & \text{if} \quad \gamma = 0, \\ \left(\dfrac{(\bar{\beta}\gamma + t)^{2m} - t^{2m}}{\bar{\beta}\gamma(\bar{\beta}\gamma + 2t)} \right)^{1/2} & \text{if} \quad \gamma > 0. \end{cases}$$

LEMMA 2.1. *If* $\mathbf{T}$ *is invertible, then there exist* $\mu \in \mathrm{sp}(W)$ *and* $\lambda \in \mathrm{sp}(A, B)$ *satisfying*

$$\|\mathbf{D}^{-1}\|_F \leqslant \frac{\varphi_\gamma(\underline{\beta}|\lambda - \mu|)}{(\underline{\beta}|\lambda - \mu|)^m}.$$

Proof. Set

$$d = \min\{|\alpha_i - l_{jj}\beta_i| : 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m\},$$

and let i_0 and j_0 be such that

$$d = |\alpha_{i_0} - l_{j_0 j_0}\beta_{i_0}|.$$

We define

$$\lambda = \alpha_{i_0}/\beta_{i_0} \quad \text{and} \quad \mu = l_{j_0 j_0}.$$

Then

$$\lambda \in \mathrm{sp}(A, B) \quad \text{and} \quad \mu \in \mathrm{sp}(W).$$

It suffices to prove the bound for $\|\tilde{\mathbf{D}}^{-1}\|_F$. Let $v = [v_1, \ldots, v_m] \in \mathbb{C}^{n \times m}$ be such that $\|v\|_F = 1$. Then $\|v_j\|_2 \leqslant 1$ for $j = 1, \ldots, m$. Let $u = [u_1, \ldots, u_m] \in \mathbb{C}^{n \times m}$ be such that $\tilde{\mathbf{D}}u = v$. It can be easily shown, by finite induction, that

$$\|u_{m-j}\|_2 \leqslant \frac{1}{d}\left(1 + \frac{\bar{\beta}\gamma}{d}\right)^j, \qquad j = 0, \ldots, m-1.$$

Hence

$$\|u\|_F^2 \leqslant \frac{1}{d^2} \sum_{j=0}^{m-1} \left(1 + \frac{\bar{\beta}\gamma}{d}\right)^{2j}.$$

It follows that

$$\|u\|_F^2 \leqslant \begin{cases} \dfrac{m}{d^2} & \text{if} \quad \gamma = 0, \\ \dfrac{1}{\bar{\beta}\gamma(\bar{\beta}\gamma + 2d)} \left[\left(1 + \dfrac{\bar{\beta}\gamma}{d}\right)^{2m} - 1\right] & \text{if} \quad \gamma > 0. \end{cases}$$

The desired bound is a consequence of the inequality

$$d \geqslant \underline{\beta}|\lambda - \mu|. \qquad \blacksquare$$

THEOREM 2.2. *Let p be the integer defined by*

$$p = \min\left\{l \in \mathbb{N}^* : (\mathbf{D}^{-1}\mathbf{N})^l = \mathbf{O}\right\}.$$

Then there exist $\mu \in \operatorname{sp}(W)$ and $\lambda \in \operatorname{sp}(A, B)$ such that for all $X \in \mathbb{C}^{n \times m}$ with norm $\|X\|_F = 1$, the following bound holds:

$$|\lambda - \mu|^{mp} \leqslant \|V\|_2 \|U^{-1}(AX - BXW)\|_F \times \sum_{k=0}^{p-1} \frac{|\lambda - \mu|^{m(p-k-1)}}{\underline{\beta}^{m(k+1)}} \left[\varphi_\gamma\left(\underline{\beta}|\lambda - \mu|\right)\right]^{k+1} \|\mathbf{N}\|_F^k,$$

with the convention that $0^0 = 1$.

Proof. If $\mathbf{T}$ is singular, then by Property 1.6, we can let $\lambda = \mu \in \operatorname{sp}(A, B) \cap \operatorname{sp}(W)$. Assume now that $\mathbf{T}$ is invertible. By Property 1.7(iii), we have

$$1 = \|X\|_F = \left\| V\left[(\mathbf{I} + \mathbf{D}^{-1}\mathbf{N})^{-1}\mathbf{D}^{-1}\right]\left[U^{-1}(AX - BXW)\right]\right\|_F.$$

But

$$(\mathbf{I} + \mathbf{D}^{-1}\mathbf{N})^{-1} = \sum_{k=0}^{p-1} (-1)^k (\mathbf{D}^{-1}\mathbf{N})^k,$$

so that

$$1 \leqslant \|V\|_2 \|U^{-1}(AX - BXW)\|_F \sum_{k=0}^{p-1} \|\mathbf{D}^{-1}\|_F^{k+1} \|\mathbf{N}\|_F^k.$$

The desired bound follows from the preceding lemma. ∎

3. SOME PARTICULAR CASES

Given a complex matrix C, we denote by $|C|$ the matrix whose entries are the absolute values of the corresponding entries of C.

The following corollaries of Theorem 2.2 deal with the case $m = 1$. Then $W = (\mu)$, $\mathbf{N}$ can be identified with $N_A - \mu N_B$, and $\mathbf{D}$ can be identified with $D_A - \mu D_B$ which is a diagonal matrix.

COROLLARY 3.1. *Given $\mu \in \mathbb{C}$, let p be the smallest positive integer such that*

$$|N_A - \mu N_B|^p = O.$$

Then there exists $\lambda \in \operatorname{sp}(A, B)$ such that for all $x \in \mathbb{C}^n$ satisfying $\|x\|_2 = 1$,

$$|\lambda - \mu|^p \leqslant \|V\|_2 \|U^{-1}(Ax - \mu Bx)\|_2 \sum_{k=0}^{p-1} \frac{|\lambda - \mu|^{p-k-1}}{\underline{\beta}^{k+1}} \|N_A - \mu N_B\|_2^k.$$

Proof. We remark that if $|N_A - \mu N_B|^l = O$ for a positive integer l, then $[D(N_A - \mu N_B)]^l = O$ for every diagonal matrix D. ∎

COROLLARY 3.2. *Suppose that B is Hermitian and positive definite, so that the matrices U and V in Theorem 2.2 can be chosen to satisfy*

$$U^{-1} = V^* \quad \textit{and} \quad V^*BV = I.$$

Let p be the smallest positive integer such that

$$|N_A|^p = O.$$

For each $\mu \in \mathbb{C}$, there exists $\lambda \in \mathrm{sp}(A, B)$ such that for all $x \in \mathbb{C}^n$ satisfying $\|x\|_2 = 1$,

$$|\lambda - \mu|^p \leqslant \|V\|_2 \|V^*(Ax - \mu Bx)\|_2 \sum_{k=0}^{p-1} |\lambda - \mu|^{p-k-1} \|N_A\|_2^k.$$

Proof. Evident, since $D_B = I$, $\underline{\beta} = 1$, and $N_B = O$. ∎

If $m = 1$ and $B = I$, that is, in the case of the ordinary eigenvalue problem, we obtain

COROLLARY 3.3. *Let U be an invertible matrix triangularizing A, that is $V = U$ in Property 1.5. Let p be the smallest positive integer such that*

$$|N_A|^p = O.$$

For each $\mu \in \mathbb{C}$, there exists $\lambda \in \mathrm{sp}(A)$ such that for all $x \in \mathbb{C}^n$ satisfying $\|x\|_2 = 1$,

$$|\lambda - \mu|^p \leqslant \|U\|_2 \|U^{-1}(Ax - \mu x)\|_2 \sum_{k=0}^{p-1} |\lambda - \mu|^{p-k-1} \|N_A\|_2^k.$$

Proof. Evident from Corollary 3.2 ∎

We remark that if A is diagonalized by U, then $p = 1$, since $N_A = O$. Then the Bauer-Householder [4] bound is obtained:

$$|\lambda - \mu| \leqslant \kappa_2(U) \|Ax - \mu x\|_2, \tag{3.1}$$

where $\kappa_2(U) = \|U\|_2 \|U^{-1}\|_2$ denotes the condition number of U relative to inversion in the $\|\cdot\|_2$ norm. When U is unitary, then $\kappa_2(U) = 1$ and the Krylov-Weinstein bound follows.

The Bauer-Fike [3] result follows if we take (μ, x) as an exact eigenpair of the perturbed matrix $A + E$, since then $Ax + \mu x = -Ex$ and hence

$$|\lambda - \mu| \leqslant \kappa_2(U)\|E\|_2, \tag{3.2}$$

when A is diagonalized by U.

More generally,

Corollary 3.4. *Let U be an invertible matrix triangularizing A, that is, $V = U$ in Property 1.5. Let p be the smallest positive integer such that*

$$|N_A|^p = O.$$

Let $\mu \in \mathbb{C}$ be an eigenvalue of $A + E$. Then there exists $\lambda \in \mathrm{sp}(A)$ such that

$$|\lambda - \mu|^p \leqslant \|U\|_2 \|U^{-1}E\|_2 \sum_{k=0}^{p-1} |\lambda - \mu|^{p-k-1} \|N_A\|_2^k.$$

Proof. Take x in Corollary 3.3 as an eigenvector of $A + E$ associated with μ and normalized by $\|x\|_2 = 1$. ∎

Corollary 3.5. *Let U be an invertible matrix such that $J = U^{-1}AU$ is a Jordan canonical form. Let p be the order of the largest Jordan block of J. If $\mu \in \mathbb{C}$ is any eigenvalue of $A + E$, then there exists $\lambda \in \mathrm{sp}(A)$ such that*

$$|\lambda - \mu|^p \leqslant \|U\|_2 \|U^{-1}E\|_2 \sum_{k=0}^{p-1} |\lambda - \mu|^k.$$

Proof. Since $J = U^{-1}AU = D_A + N_A$ is a Jordan canonical form, $\|N_A\|_2 \leqslant 1$ and $|N_A|^p = O$. If μ an eigenvalue of $A + E$, then we take x in Corollary 3.3 as an eigenvector of $A + E$ associated with μ and normalized by $\|x\|_2 = 1$. ∎

Ahues's result [1] follows readily: If $\mathrm{sp}(A) = \{\lambda\}$ and l is the ascent of λ, then for any eigenvalue μ of $A + E$,

$$|\lambda - \mu|^l \leqslant \kappa_2(U)\|E\|_2 \sum_{k=0}^{l-1} |\lambda - \mu|^k. \tag{3.3}$$

Similarly, we deduce the bound given by Golub and Van Loan [10]:

$$|\lambda - \mu| \leqslant \max\{\varepsilon, \varepsilon^{1/p}\}, \tag{3.4}$$

where

$$\varepsilon = \|E\|_2 \sum_{k=0}^{p-1} \|N_A\|_2^k,$$

μ being an eigenvalue, of $A + E$, N_A the strictly upper triangular part of a Schur triangularization of A (that is, $V^* = U^* = U^{-1}$ in Property 1.5), and p the smallest positive integer such that $|N_A|^p = O$.

If the columns of U form an ordered basis of $\mathbb{C}^n$, composed of ordered bases of various maximal invariant subspaces of A, then $U^{-1}AU$ is a block-diagonal matrix. The columns of U may be chosen in such a way that each diagonal block of $U^{-1}AU$ is an upper triangular matrix. Hence $U^{-1}AU$ is upper triangular: $U^{-1}AU = D_A + N_A$, with N_A block-diagonal and strictly upper triangular. The bound given by Chu [7] follows easily:

$$|\lambda - \mu| \leqslant \max\{\varepsilon, \varepsilon^{1/p}\}, \tag{3.5}$$

where

$$\varepsilon = \kappa_2(U)\|E\|_2\|N_A\|_2,$$

μ being an eigenvalue of $A + E$ and p the order of the largest block on the diagonal of $U^{-1}AU$.

The following bound is due to De Boor and Swartz [8]:

$$|\lambda - \mu|^p \leqslant \|E\|_2 \sum_{k=0}^{p-1} \left\|(A + E - \lambda I)^{p-k-1}\right\|_2 \left\|(A - \lambda I)^k\right\|_2, \tag{3.6}$$

where $(A - \lambda I)^p = O$ and $\mu \in \operatorname{sp}(A + E)$. Our result improves on (3.6), since $(\mu - \lambda)^{p-k-1}$ is an eigenvalue of $(A + E - \lambda I)^{p-k-1}$ and hence

$$\left|(\mu - \lambda)^{p-k-1}\right| \leqslant \left\|(A + E - \lambda I)^{p-k-1}\right\|_2.$$

Also, two results given by Henrici [12, 15] can be obtained from our main theorem. Let $s \in \mathbb{R}_+$ be given and let $g(s)$ be the unique nonnegative solution of the equation

$$\sum_{k=1}^{n} \xi^k = s.$$

Set

$$\mathbf{n}(A) = \inf\{\|N\|_2 : U^*AU = D + N,\ U \text{ unitary},$$
$$D \text{ diagonal, } N \text{ strictly upper triangular}\}.$$

If $\mu \in \mathbb{C}$ and $x \in \mathbb{C}^n$ are such that $\|x\|_2 = 1$ and $\|Ax - \mu x\|_2 \neq 0$, then there exists $\lambda \in \mathrm{sp}(A)$ such that

$$|\lambda - \mu| \leqslant \frac{\mathbf{n}(A)}{g\left(\dfrac{\mathbf{n}(A)}{\|Ax - \mu x\|_2}\right)}, \tag{3.7}$$

which is Morrison's result quoted by Henrici (Theorem 6 in [12]). Further, if μ is an eigenvalue of $A + E$ and $E \neq O$, then (cf. Theorem 4 in [12])

$$|\lambda - \mu| \leqslant \frac{\mathbf{n}(A)}{g\left(\dfrac{\mathbf{n}(A)}{\|E\|_2}\right)}. \tag{3.8}$$

To deduce these bounds from our results, it suffices to consider a unitary matrix U in Corollary 3.3 and take the infimum of $\|N_A\|_2$ over all unitary matrices U triangularizing A. This leads to

$$|\lambda - \mu|^p \leqslant \|(Ax - \mu x)\|_2 \sum_{k=0}^{p-1} |\lambda - \mu|^{p-k-1} \mathbf{n}(A)^k,$$

which is equivalent to (3.7). In a similar manner, Corollary 3.4 leads to (3.8).

Both (3.7) and (3.8) can be extended to the generalized eigenvalue problem. We first remark that, using QR factorization, the matrices U and V of Property 1.5 may be chosen to be unitary. Hence we can define

$$\mathbf{n}(\mu, A, B) = \inf\{\|D_B^{-1}\|_2 \|N_A - \mu N_B\|_2 :$$
$$U^*AV = D_A + N_A,\ U^*BV = D_B + N_B,$$

$$U, V \text{ unitary}, D_A, D_B \text{ diagonal},$$

$$N_A, N_B \text{ strictly upper triangular}\}$$

and

$$\hat{\beta} = \inf\left\{\|D_B^{-1}\|_2^{-1} : U^*BV = D_B + N_B,\ D_B \text{ diagonal}, \right.$$

$$\left. N_B \text{ strictly upper triangular},\ U, V \text{ unitary}\right\}.$$

Let U and V be unitary matrices such that

$$U^*BV = D_B + N_B,$$

with $D_B = \operatorname{diag}(\beta_1, \dots, \beta_n)$ and N_B strictly upper triangular. Let i be such that $\|D_B^{-1}\|_2^{-1} = |\beta_i|$. Then

$$\|B^{-1}\|_2 = \left\|(D_B + N_B)^{-1}\right\| \geqslant \left\|e_i e_i^*(D_B + N_B)^{-1} e_i e_i^*\right\|_2 = \frac{1}{|\beta_i|}.$$

Hence

$$\hat{\beta} \geqslant \frac{1}{\|B^{-1}\|_2}.$$

We obtain the following generalizations. If $\mu \in \mathbb{C}$ and $x \in \mathbb{C}^n$ are such that $\|x\|_2 = 1$ and $\|(A - \mu B)x\|_2 \neq 0$, then there exists $\lambda \in \operatorname{sp}(A, B)$ such that

$$|\lambda - \mu| \leqslant \frac{\mathbf{n}(\mu, A, B)}{g\left(\dfrac{\hat{\beta}\mathbf{n}(\mu, A, B)}{\|(A - \mu B)x\|_2}\right)}. \tag{3.9}$$

Also, when μ is an eigenvalue of the pencil $(A + E, B + F)$ and $E - \mu F \neq O$, then

$$|\lambda - \mu| \leqslant \frac{\mathbf{n}(\mu, A, B)}{g\left(\dfrac{\hat{\beta}\mathbf{n}(\mu, A, B)}{\|E - \mu F\|_2}\right)}. \tag{3.10}$$

Finally, we compare a result given in Theorem VI 2.6 of [18] with a bound that can be deduced from Corollary 3.1. In our notation the quoted result reads as follows:

Let (A, B) *be a regular pencil such that there exist invertible matrices* U *and* V *diagonalizing* (A, B):

$$U^{-1}AV = D_A, \qquad U^{-1}BV = D_B.$$

Let $(A + E,\ B + F)$ *be the perturbed pencil which is assumed to be regular. Then, for any* $\mu \in \mathrm{sp}(A + E,\ B + F)$, *there exists* $\lambda \in \mathrm{sp}(A, B)$ *such that*

$$\chi(\lambda, \mu) \leqslant \kappa_2(V)\,\rho_L[(A, B), (A + E, B + F)], \tag{3.11}$$

where

$$\chi(\lambda, \mu) = \frac{|\lambda - \mu|}{\sqrt{1 + |\lambda|^2}\sqrt{1 + |\mu|^2}}$$

is the chordal metric,

$$\rho_L[(A, B), (A + E,\ B + F)] = \sin \theta_{\max},$$

and $\theta_{\max}$ *is the largest canonical angle between the row spaces of the* $n \times 2n$ *matrices* $(A\ \ B)$ *and* $(A + E\ \ B + F)$.

Corollary 3.1 allows us to find a similar bound as follows. Since B is invertible, $H = AA^* + BB^*$ is a Hermitian definite positive matrix. Let $H^{1/2}$ denote its unique Hermitian positive definite square root. We set

$$\hat{A} = H^{-1/2}A \quad \text{and} \quad \hat{B} = H^{-1/2}B.$$

Then $[\hat{A}\ \hat{B}]$ has orthonormal rows and $\mathrm{sp}(\hat{A}, \hat{B}) = \mathrm{sp}(A, B)$. Moreover, if (U, V) diagonalizes (A, B) in the sense of Property 1.5, then $(\hat{A}, \hat{B})$ is diagonalized by $(H^{-1/2}U, V)$. We remark that

$$\|H^{1/2}\|_2 = \rho(H^{1/2}) = [\rho(H)]^{1/2} = \|[A \quad B]\|_2.$$

We apply Corollary 3.1 to the pencil $(\hat{A}, \hat{B})$ in the particular case of $p = 1$ and $N_{\hat{A}} = N_{\hat{B}} = O$. Then, for any $\mu \in \mathbb{C}$, there exists $\lambda \in \mathrm{sp}(A, B)$ such that for any unit vector x,

$$\chi(\lambda, \mu) \leqslant \frac{|\lambda - \mu|}{\sqrt{1 + |\mu|^2}} \leqslant \|V\|_2 \frac{\left\|U^{-1}H^{1/2}(\hat{A} - \mu\hat{B})x\right\|_2}{\underline{\beta}\sqrt{1 + |\mu|^2}}$$

$$\leqslant \frac{\|H^{1/2}\|_2}{\underline{\beta}} \|U^{-1}\|_2 \|V\|_2 \frac{\left\|(\hat{A} - \mu\hat{B})x\right\|_2}{\sqrt{1 + |\mu|^2}}.$$

In a similar way, if

$$G = (A + E)(A + E)^* + (B + F)(B + F)^*,$$

$$\widehat{\underline{A} + \underline{E}} = G^{-1/2}(A + E), \qquad \widehat{B + F} = G^{-1/2}(B + F),$$

then $[\widehat{A + E}\ \ \widehat{B + F}]$ has orthonormal rows and $\mathrm{sp}(A + E, B + F) = \mathrm{sp}(\widehat{A + E}, \widehat{B + F})$. If $\mu \in \mathrm{sp}(A + E, B + F)$ and if x is a unit eigenvector of $(\widehat{A + E}, \widehat{B + F})$ corresponding to μ, then, as is proved in [18],

$$\frac{\left\|(\hat{A} - \mu\hat{B})x\right\|_2}{\sqrt{1 + |\mu|^2}} \leqslant \rho_L\left[(\hat{A}, \hat{B}), (\widehat{A + E}, \widehat{B + F})\right].$$

But

$$\rho_L\left[(\hat{A}, \hat{B}), (\widehat{A + E}, \widehat{B + F})\right] = \rho_L[(A, B), (A + E, B + F)].$$

Let us define

$$\mathscr{M} = \Big\{(U, V) \in \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n} : U, V \text{ are invertible,}$$
$$U^{-1}AV \text{ and } U^{-1}BV \text{ are diagonal and}$$
$$\|V^{-1}B^{-1}U\|_2^{-1} = \|[A \quad B]\|_2\Big\}$$

and

$$\kappa = \inf_{(U, V) \in \mathscr{M}} \|U^{-1}\|_2 \|V\|_2.$$

We deduce that

If B is invertible, then for all $\mu \in \mathrm{sp}(A+E, B+F)$ there exists $\lambda \in \mathrm{sp}(A, B)$ such that

$$\chi(\lambda, \mu) \leqslant \kappa \rho_L[(A, B), (A+E, B+F)].$$

4. FINAL REMARKS

4.1

In the following example, only one eigenvalue of W satisfies the bound in Theorem 2.2: We set $n = m = 2$, $W = \mathrm{diag}(\mu, \mu_0)$, $A = \lambda I_2$, $B = I_2$, so that $\gamma = 0$, $\varphi_\gamma(t) = t\sqrt{2}$, $N_A = N_B = O$, $\mathbf{N} = \mathbf{O}$, $\underline{\beta} = 1$, $U = V = I_2$, $p = 1$, $D_A = A$, and $D_B = B$.

Let us choose λ, μ, and μ_0 such that

$$0 < \sqrt{2}|\lambda - \mu| < |\lambda - \mu_0|.$$

Then the pair $(\lambda, \mu) \in \mathrm{sp}(A, B) \times \mathrm{sp}(W)$ satisfies the bound in Theorem 2.2 for any $x \in \mathbb{C}^2$. However, with

$$x = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbb{C}^2,$$

the pair (λ, μ_0) in $\mathrm{sp}(A, B) \times \mathrm{sp}(W)$ does not verify the bound.

4.2

Different definitions of the *departure from normality of a matrix W* have been given in the literature (cf. [6, 10, 13]). We consider here the most direct one:

$$\nu_F(W) = \|W^*W - WW^*\|_F,$$

which is obviously zero if and only if W is a normal matrix.

The quantity γ introduced at the beginning of Section 2 is related to the departure from normality of W. In fact (cf. [12]),

$$\frac{\nu_F(W)^2}{6\|W\|_F^2} \leqslant \|L\|_F^2 \leqslant \sqrt{\frac{m^3 - m}{12}}\, \nu_F(W),$$

but

$$\gamma^2 \leqslant \|L\|_F^2 \leqslant \frac{m(m-1)}{2}\gamma^2,$$

so that

$$\frac{\nu_F(W)^2}{3m(m-1)\|W\|_F^2} \leqslant \gamma^2 \leqslant \sqrt{\frac{m^3-m}{12}}\,\nu_F(W).$$

4.3

Kahan, Parlett, and Jiang [13] propose the bound

$$|\lambda - \mu|^l \leqslant \|UEU^{-1}\|_2(1 + |\lambda - \mu|)^{l-1}, \tag{4.1}$$

where U is a Jordan basis of A and l is the order of the largest Jordan block to which λ belongs. This result should be compared with our bound in Corollary 3.5.

If $p = l$ and $\|UEU^{-1}\|_2 = \|U\|_2\|U^{-1}E\|_2$, then our result gives an upper bound of $|\lambda - \mu|^p$, which is sharper than (4.1). This situation occurs, for instance, if n and p satisfy $2(p-1) > n \geqslant p \geqslant 3$ and we take

$$A = \lambda I_n + N, \qquad N = \begin{bmatrix} 0 & I_{p-1} \\ 0 & 0 \end{bmatrix}, \quad \text{and} \quad E = \epsilon I_n,$$

since then $U = I_n$, $l = p$, and Corollary 3.5 gives

$$|\lambda - \mu|^p < \sum_{k=0}^{p-1} \epsilon^k,$$

while (4.1) gives

$$|\lambda - \mu|^p \leqslant \sum_{k=0}^{p-1} \binom{p-1}{k} \epsilon^k.$$

4.4

In the bounds (3.9) and (3.10) we remark that if $B = I$, then $\hat{\beta} = 1$. This shows that these bounds actually generalize the bounds (3.7) and (3.8) respectively.

4.5

The use of the Euclidean vector norm $\|\cdot\|_2$ and the corresponding subordinated matrix norm is not essential, except for the fact that $\|U\|_2 = 1$ for any unitary matrix U. Almost all of the results presented in this work can be demonstrated using a matrix submultiplicative norm $\|\cdot\|_\star$ such that for any diagonal matrix $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ one has

$$\|D\|_\star \leqslant \max_{1 \leqslant i \leqslant n} |\lambda_i|,$$

together with a vector norm $\|\cdot\|$ satisfying

$$\|Mx\| \leqslant \|M\|_\star \|x\|$$

for all $M \in \mathbb{C}^{n \times n}$ and $x \in \mathbb{C}^n$.

Thanks are due to Rafikul Alam and Alain Largillier as well as to the referee of a previous version for helpful comments and suggestions.

REFERENCES

1 M. Ahues, Spectral condition numbers for defective eigenelements of linear operators in Hilbert spaces, *Numer. Funct. Anal. Optim.* 10(9 & 10): 843–861 (1989).
2 M. Ahues and F. Chatelin, *Exercises de Valeurs Propres de Matrices*, Collect. Math. Appl. Maîtrise, Masson, Paris, 1989.
3 F. L. Bauer and C. T. Fike, Norms and exclusion theorems, *Numer. Math.* 2:137–141 (1960).
4 F. L. Bauer and A. S. Householder, Moments and characteristic roots, *Numer. Math.* 2:42–53 (1960).
5 F. L. Bauer and A. S. Householder, Absolute norms and characteristic roots, *Numer. Math.* 3:241–246 (1961).
6 F. Chatelin, *Valeurs Propres de Matrices*, Collect. Math. Appl. Maîtrise, Masson, Paris, 1988.
7 K. E. Chu, Generalization of the Bauer-Fike theorem, *Numer. Math.* 49:685–691 (1986).
8 C. De Boor and B. Swartz, Collocation approximation to eigenvalues of an ordinary differential equation: The principle of the thing, *Math. Comp.* 35:679–694 (1980).
9 I. Gohberg, S. Goldberg, and M. A. Kaashoek, *Classes of Linear Operators*, Vol. I.
10 G. H. Golub and C. Van Loan, *Matrix Computations*, Johns Hopkins U.P., Baltimore, 1989.

11 A. Graham, *Kronecker Products and Matrix Calculus with Applications*, Ellis Horwood, Chichester, 1981.
12 P. Henrici, Bounds for iterates, inverses, spectral variation and fields of values of the spectrum of nonnormal matrices, *Numer. Math.* 4:24–40 (1962).
13 W. M. Kahan, B. N. Parlett, and E. Jiang, Residual bounds on approximate eigensystems of nonnormal matrices, *SIAM J. Numer. Anal.* 19:470–484 (1982).
14 P. Lancaster, Explicit solutions of linear matrix equations, *SIAM Rev.* 12:544–566 (1970).
15 D. D. Morrison, Ph.D. Dissertation, Univ. of California, Los Angeles, 1961.
16 G. W. Stewart, On the sensitivity of the eigenvalue problem $Ax = \lambda Bx$, *SIAM J. Numer. Anal.* 9:669–686 (1972).
17 G. W. Stewart, Error and perturbation bounds for subspaces associated with certain eigenvalue problems, *SIAM Rev.* 15:727–764 (1973).
18 G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory*, Academic, San Diego, 1990.
19 J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford U.P., 1965.

Received 19 May 1994; final manuscript accepted 29 January 1997